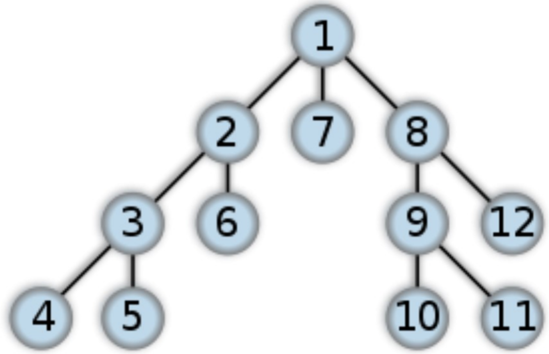
Depth-first search

From Wikipedia, the free encyclopedia

Depth-first search (**DFS**) is an algorithm for traversing or searching tree or graph data structures. One starts at the root (selecting some arbitrary node as the root in the case of a graph) and explores as far as possible along each branch before backtracking.

A version of depth-first search was investigated in the 19th century by French mathematician Charles Pierre Trémaux[1] as a strategy for solving mazes.[2][3]

Depth-first search



Order in which the nodes are visited

Class	Search algorithm
Data structure	Graph
Worst-case performance	$O(\|V\|+\|E\|)$ for explicit graphs traversed without repetition, $O(b^d)$ for implicit graphs with branching factor *b* searched to depth *d*
Worst-case space complexity	$O(\|V\|)$ if entire graph is traversed without repetition, O(longest path length searched) for implicit graphs without elimination of duplicate nodes

Contents

- 1 Properties
- 2 Example
- 3 Output of a depth-first search
  - 3.1 DFS ordering
  - 3.2 Vertex orderings
- 4 Pseudocode
- 5 Applications
- 6 Complexity
- 7 See also
- 8 Notes
- 9 References
- 10 External links

Properties

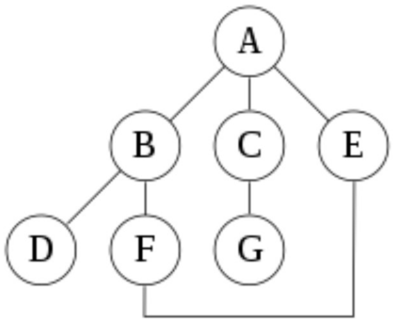
The time and space analysis of DFS differs according to its application area. In theoretical computer science, DFS is typically used to traverse an entire graph, and takes time $\Theta(|V| + |E|)$,[4] linear in the size of the graph. In these applications it also uses space $O(|V|)$ in the worst case to store the stack of vertices on the current search path as well as the set of already-visited vertices. Thus, in this setting, the time and space bounds are the same as for breadth-first search and the choice of which of these two algorithms to use depends less on their complexity and more on the different properties of the vertex orderings the two algorithms produce.

For applications of DFS in relation to specific domains, such as searching for solutions in artificial intelligence or web-crawling, the graph to be traversed is often either too large to visit in its entirety or infinite (DFS may suffer from non-termination). In such cases, search is only performed to a limited depth; due to limited resources, such as memory or disk space, one typically does not use data structures to keep track of the set of all previously visited vertices. When search is performed to a limited depth, the time is still linear in terms of the number of expanded vertices and edges (although this number is not the same as the size of the entire graph because some vertices may be searched more than once and others not at all) but the space complexity of this variant of DFS is only proportional to the depth limit, and as a result, is much smaller than the space needed for searching to the same depth using breadth-first search. For such applications, DFS also lends itself much better to heuristic methods for choosing a likely-looking branch. When an appropriate depth limit is not known a priori, iterative deepening depth-first search applies DFS repeatedly with a sequence of increasing limits. In the artificial intelligence mode of analysis, with a branching factor greater than one, iterative deepening increases the running time by only a constant factor over the case in which the correct depth limit is known due to the geometric growth of the number of nodes per level.

DFS may also be used to collect a sample of graph nodes. However, incomplete DFS, similarly to incomplete BFS, is biased towards nodes of high degree.

Example

For the following graph:

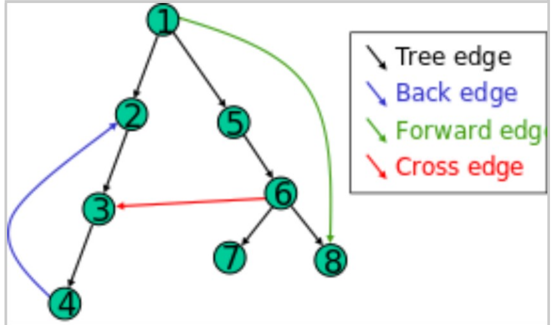


a depth-first search starting at A, assuming that the left edges in the shown graph are chosen before right edges, and assuming the search remembers previously visited nodes and will not repeat them (since this is a small graph), will visit the nodes in the following order: A, B, D, F, E, C, G. The edges traversed in this search form a Trémaux tree, a structure with important applications in graph theory.

Performing the same search without remembering previously visited nodes results in visiting nodes in the order A, B, D, F, E, A, B, D, F, E, etc. forever, caught in the A, B, D, F, E cycle and never reaching C or G.

Iterative deepening is one technique to avoid this infinite loop and would reach all nodes.

Output of a depth-first search



The four types of edges defined by a spanning tree

A convenient description of a depth first search of a graph is in terms of a spanning tree of the vertices reached during the search. Based on this spanning tree, the edges of the original graph can be divided into three classes: **forward edges**, which point from a node of the tree to one of its descendants, **back edges**, which point from a node to one of its ancestors, and **cross edges**, which do neither. Sometimes **tree edges**, edges which belong to the spanning tree itself, are classified separately from forward edges. If the original graph is undirected then all of its edges are tree edges or back edges.

DFS ordering

An enumeration of the vertices of a graph is said to be a DFS ordering if it is the possible output of the application of DFS to this graph.

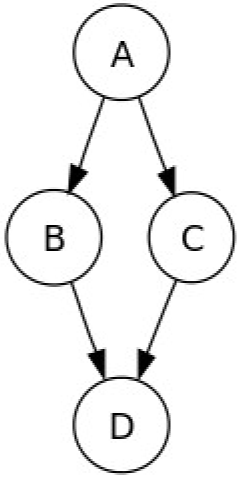
Let $G = (V, E)$ be a graph with n vertices. For $\sigma = (v_1, \ldots, v_m)$ be a list of distinct elements of V, for $v \in V \setminus \{v_1, \ldots, v_m\}$, let $\nu_\sigma(v)$ be the greatest i such that v_i is a neighbor of v, if such a i exists, and be 0 otherwise.

Let $\sigma = (v_1, \ldots, v_n)$ be an enumeration of the vertices of V. The enumeration σ is said to be a DFS ordering (with source v_1) if, for all $1 < i \leq n$, v_i is the vertex $w \in V \setminus \{v_1, \ldots, v_i - 1\}$ such that $\nu_{(v_1, \ldots, v_{i-1})}(w)$ is maximal. Recall that $N(v)$ is the set of neighbors of v. Equivalently, σ is a DFS ordering if, for all $1 \leq i < j < k \leq n$ with $v_i \in N(v_k) \setminus N(v_j)$, there exists a neighbor v_m of v_j such that $i < m < j$.

Vertex orderings

It is also possible to use the depth-first search to linearly order the vertices of the original graph (or tree). There are three common ways of doing this:

- A **preordering** is a list of the vertices in the order that they were first visited by the depth-first search algorithm. This is a compact and natural way of describing the progress of the search, as was done earlier in this article. A preordering of an expression tree is the expression in Polish notation.
- A **postordering** is a list of the vertices in the order that they were *last* visited by the algorithm. A postordering of an expression tree is the expression in reverse Polish notation.
- A **reverse postordering** is the reverse of a postordering, i.e. a list of the vertices in the opposite order of their last visit. Reverse postordering is not the same as preordering. For example, when searching the directed graph in pre-order



beginning at node A, one visits the nodes in sequence, to produce lists either A B D B A C A, or A C D C A B A (depending upon whether the algorithm chooses to visit B or C first). Note that repeat visits in the form of backtracking to a node, to check if it has still unvisited neighbours, are included here (even if it is found to have none). Thus the possible preorderings are A B D C and A C D B (order by node's leftmost occurrence in above list), while the possible reverse postorderings are A C B D and A B C D (order by node's rightmost occurrence in above list). Reverse postordering produces a topological sorting of any directed acyclic graph. (Possible postordering are D B C A and D C B A.) This ordering is also useful in control flow analysis as it often represents a natural linearization of the control flows. The graph above might represent the flow of control in a code fragment like

```
if (A) then {
  B
} else {
  C
}
D
```

and it is natural to consider this code in the order A B C D or A C B D, but not natural to use the order A B D C or A C D B.

Pseudocode

Input: A graph *G* and a vertex *v* of G

Output: All vertices reachable from *v* labeled as discovered

A recursive implementation of DFS:[5]

```
procedure DFS(G,v):
    label v as discovered
    for all edges from v to w in G.adjacentEdges(v) do
        if vertex w is not labeled as discovered then
            recursively call DFS(G,w)
```

A non-recursive implementation of DFS:[6]

```
procedure DFS-iterative(G,v):
    let S be a stack
    S.push(v)
    while S is not empty
        v = S.pop()
        if v is not labeled as discovered:
            label v as discovered
            for all edges from v to w in G.adjacentEdges(v) do
                S.push(w)
```

These two variations of DFS visit the neighbors of each vertex in the opposite order from each other: the first neighbor of *v* visited by the recursive variation is the first one in the list of adjacent edges, while in the iterative variation the first visited neighbor is the last one in the list of adjacent edges. The recursive implementation will visit the nodes from the example graph in the following order: A, B, D, F, E, C, G. The non-recursive implementation will visit the nodes as: A, E, F, B, D, C, G.

The non-recursive implementation is similar to breadth-first search but differs from it in two ways:

1. it uses a stack instead of a queue, and
2. it delays checking whether a vertex has been discovered until the vertex is popped from the stack rather than making this check before pushing the vertex.

Note that this non-recursive implementation of DFS may use O(|E|) space in the worst case, for example on a complete graph.

Applications

Algorithms that use depth-first search as a building block include:

- Finding connected components.
- Topological sorting.
- Finding 2-(edge or vertex)-connected components.

- Finding 3-(edge or vertex)-connected components.
- Finding the bridges of a graph.
- Generating words in order to plot the Limit Set of a Group.
- Finding strongly connected components.
- Planarity testing[7][8]
- Solving puzzles with only one solution, such as mazes. (DFS can be adapted to find all solutions to a maze by only including nodes on the current path in the visited set.)
- Maze generation may use a randomized depth-first search.
- Finding biconnectivity in graphs.

Randomized algorithm similar to depth-first search used in generating a maze.

Complexity

The computational complexity of DFS was investigated by John Reif, who showed that a decision version of it (establish whether some vertex u occurs before some vertex v in a DFS order of a rooted graph) is **P**-complete,[9] meaning that it is "a nightmare for parallel processing".[10]:189

See also

- Tree traversal (for details about pre-order, in-order and post-order depth-first traversal)
- Breadth-first search
- Iterative deepening depth-first search
- Search games

Notes

1. Charles Pierre Trémaux (1859–1882) École polytechnique of Paris (X:1876), French engineer of the telegraph in Public conference, December 2, 2010 – by professor Jean Pelletier-Thibert in Académie de Macon (Burgundy – France) – (Abstract published in the Annals academic, March 2011 – ISSN 0980-6032 (https://www.worldcat.org/search?fq=x0:jrnl&q=n2:0980-6032))
2. Even, Shimon (2011), *Graph Algorithms* (2nd ed.), Cambridge University Press, pp. 46–48, ISBN 978-0-521-73653-4.
3. Sedgewick, Robert (2002), *Algorithms in C++: Graph Algorithms* (3rd ed.), Pearson Education, ISBN 978-0-201-36118-6.
4. Cormen, Thomas H., Charles E. Leiserson, and Ronald L. Rivest. p.606
5. Goodrich and Tamassia; Cormen, Leiserson, Rivest, and Stein
6. Page 93, Algorithm Design, Kleinberg and Tardos
7. Hopcroft, John; Tarjan, Robert E. (1974), "Efficient planarity testing", *Journal of the Association for Computing Machinery*, **21** (4): 549–568,

doi:10.1145/321850.321852.
8. de Fraysseix, H.; Ossona de Mendez, P.; Rosenstiehl, P. (2006), "Trémaux Trees and Planarity", *International Journal of Foundations of Computer Science*, **17** (5): 1017–1030, doi:10.1142/S0129054106004248.
9. Reif, John H. (1985). "Depth-first search is inherently sequential". *Information Processing Letters*. **20** (5). doi:10.1016/0020-0190(85)90024-9.
10. Mehlhorn, Kurt; Sanders, Peter (2008). *Algorithms and Data Structures: The Basic Toolbox* (PDF). Springer.

References

- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Section 22.3: Depth-first search, pp. 540–549.
- Goodrich, Michael T.; Tamassia, Roberto (2001), *Algorithm Design: Foundations, Analysis, and Internet Examples*, Wiley, ISBN 0-471-38365-1
- Kleinberg, Jon; Tardos, Éva (2006), *Algorithm Design*, Addison Wesley, pp. 92–94
- Knuth, Donald E. (1997), *The Art of Computer Programming Vol 1. 3rd ed*, Boston: Addison-Wesley, ISBN 0-201-89683-4, OCLC 155842391

External links

Wikimedia Commons has media related to ***Depth-first search***.

- Open Data Structures - Section 12.3.2 - Depth-First-Search (http://opendatastructures.org/versions/edition-0.1e/ods-java/12_3_Graph_Traversal.html#SECTION001532000000000000000)
- C++ Boost Graph Library: Depth-First Search (http://www.boost.org/libs/graph/doc/depth_first_search.html)
- Depth-First Search Animation (for a directed graph) (http://www.cs.duke.edu/csed/jawaa/DFSanim.html)
- Depth First and Breadth First Search: Explanation and Code (http://www.kirupa.com/developer/actionscript/depth_breadth_search.htm)
- QuickGraph (http://quickgraph.codeplex.com/Wiki/View.aspx?title=Depth%20First%20Search%20Example), depth first search example for .Net
- Depth-first search algorithm illustrated explanation (Java and C++ implementations) (http://www.algolist.net/Algorithms/Graph_algorithms/Undirected/Depth-first_search)
- YAGSBPL – A template-based C++ library for graph search and planning (https://code.google.com/p/yagsbpl/)

Categories: Graph algorithms | Search algorithms